DOCUMENT RESUME

ED 319 146 EA 021 885

AUTHOR Grubb, Ralph E.
TITLE Effects of Paired Student Interaction in the Computer Tutoring of Statistics.
PUB DATE 64
NOTE 14p.
PUB TYPE Reports - Research/Technical (143)

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS Analysis of Variance; *Computer Assisted Instruction; Cost Effectiveness; Higher Education; *Peer Teaching; *Statistics; *Tutoring
IDENTIFIERS *Paired Student Interaction

ABSTRACT

Although computer-assisted instruction (CAI) simulates a conversation between a tutor and student, it does not permit the quality of interaction so desirable in the tutorial process. This study attempted to see if team learning techniques might be one answer to the lack of interaction in programmed instruction. What would happen if pairs of students took a CAI course cooperatively? Thirty liberal arts college students were selected for this experiment on a predictor variable, College Entrance Examination Board (CEEB) scores (verbal only), to take a computer-guided statistics course. Students were divided into low- and high-scoring groups. In each group, 10 students formed pairs and 5 worked separately as controls. After the final exam, students were administered a questionnaire designed to measure their perceptions of their performance and those of their partners, if they were paired. The results from this study seem to indicate that students paired on CEEB verbal scores as a predictor variable will do as well as their controls on a final exam in a CAI course. In addition, they can complete the course in the same amount of time as their controls. The economic advantage is quickly realized since the cost of the educational terminal device has been cut by a factor of two in the process. (MLF)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED319146

THE EFFECTS OF PAIRED STUDENT INTERACTION IN THE COMPUTER TUTORING OF STATISTICS

by
Ralph E. Grubb
Behavioral Sciences Group
Thomas J. Watson Research Center
Yorktown Heights, New York
1964

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
This document has been reproduced as received from the person or organization originating it.
Minor changes have been made to improve reproduction quality.
Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

Now at IBM, Los Gatos, California

EA 021 885
BEST COPY AVAILABLE

Reprinted February 6, 1967

THE EFFECTS OF PAIRED STUDENT INTERACTION IN THE COMPUTER TUTORING OF STATISTICS

by
Ralph E. Grubb

INTRODUCTION

A topic in programed instruction that has provoked concern in educational circles is the fact that the student learns in isolation. While the classroom is potentially wealthy in dynamic social interactions, programed instruction has partly justified the exclusion of these experiences because it promised best sequences of learning. In recent years, however, some have felt that neither the classroom nor programed instruction was measuring up to these ideals.

It was in these space-time coordinates that the Behavioral Sciences group at the Thomas J. Watson Research Center began to investigate computer-assisted instruction (CAI). Because the computer was simulating a conversation taking place between a tutor and student, Uttal[4] coined the term "conversational interaction" to describe this process. The rationale was that a general purpose stored program computer could not only simulate much of the tutoring dialog but could adapt its learning sequences to individual students as well.

It soon became apparent to this investigator though, that simulated interaction in the present state of the art would not permit one to arrive at the quality of interaction so idealized in the tutorial process. Not that interaction per se is sacrosanct, but the fact that it provides the learner with the opportunity to articulate insights that ordinarily might remain pre-verbal is what makes this process so desirable. In addition, it permits the student to exchange learner-tutor roles thereby making learning a more active process.

The present study was born in this context in order to see if team learning techniques might be one answer to the lack of interaction in PI. That is to say, what would happen if pairs of students took a CAI course cooperatively? How would it affect time to complete the course, final performance measures and certain attitudes?

METHOD

Thirty liberal arts college students were selected for this experiment on a predictor variable, College Entrance Examination Board scores (verbal only)* to take the computer guided statistics course. This course covers both descriptive and inferential statistics for students in psychology and education and is taught in a "guided discovery" manner. (See Grubb and Selfridge[3] for an earlier description of the course and teaching logic.)

Since the current national mean for the CEEB is 444, two score classe were chosen for student selection: 300-400 and 500-600, which will be referred to as Low and High groups respectively.**

The following table summarizes the classification of the experimental and control Ss for the experimental design of the study.

---

* A prior unpublished pilot study showed a correlation of .74 between CEEB verbal scores and final exam performance.

** In an absolute sense, neither the H's nor the L's are so high or low as one would desire. Because E felt it desirable to have the two groups distributed somewhat symmetrically around the mean, going any higher on the upper end would force the experimenter out of the college market on the low end.

Table I

Treatments in Paired Students Learning

group	n
High pairs (HP)	10
High controls (HC)	5
Low pairs (LP)	10
Low controls (LC)	5
Total	30

Both males and females participated in the study, however the pair at the terminal was always the same sex.

While pairs were told that during the course they could converse in any way they desired to arrive at an answer, the only ground rule was they had to agree on an answer before entering it into their typewriter teaching station (they could of course agree to disagree). This procedure was intended to serve as a partial safeguard for the submissive type person that happened to be paired with a dominant individual.

Students worked at the teaching stations two hours a day, three days a week in blocks of a total of six Ss each (HP, HC, LP, LC). All students were examined individually approximately two days after completion of the course. The examination was of the paper and pencil type which consisted almost exclusively of problem solving and computational questions, i. e. test for significance between two means and accept or reject the null hypothesis at a stated level of confidence.

After the final exam, each student was administered a questionnaire designed to measure his perception of his performance and that of his partner's, if he had one. Four of the more interesting questions were as follows:

(No. 3) Assuming the examination to be worth 100 points, what would you estimate your score to be? (No. 4) Would you estimate that your partner's score was higher or lower than yours on the final exam? (No. 5) If you took another course under these conditions, would you prefer to work alone or with a partner? (No. 6) If you did work with a partner under these conditions in the future, would you prefer to work with the same partner as in this course or a different partner?

RESULTS

Instruction Time and Final Performance

Means and standard deviations for time to complete the instructional material as well as final exam performance is reported for the four treatment groups in Table II.

Table II

Mean Instructional Time and Final Performance

Treatment	Mean Time	sd	Mean Performance	sd
High Pairs	10.03 hrs.	1.00	74.9%	13.1
High Controls	11.02	.72	79.0	11.4
Low Pairs	12.27	3.09	71.0	18.4
Low Controls	12.05	2.78	69.0	12.5

While there is an apparent difference between treatments in mean time to complete the instructional material, this difference was not significant in view of the large standard deviations for LP (3.09) and LC (2.78).

The inter-quartile range in instruction time computed across all students was 2.15 hours.

The analysis of variance model was used to test for significance of difference in final exam scores between the four treatment groups. A

Cochran C test demonstrated that the assumption of homogenity of variance could not be rejected.

The analysis indicated that there is no significant difference in final exam performance between any of the treatments in this study. The apparent difference that does exist between all Highs and all Lows, however, suggests a weak trend in the expected direction ($F = 2.21$, $< .10$ $p < .25$; 1, 12 d. f.) and might merit practical considerations as well as further research.

A Pearson Product correlation coefficient computed between time to complete the course and final exam performance was -. 23. This low negative relationship indicates that there was some tendency for people requiring more time to complete the course to score lower on the final exam.

The relationship between verbal CEEB score and final exam performance is reflected in a low positive correlation of . 33.

Error Rates

A further look in depth at the learning process in this study is an analysis of the mean number of error cues issued by the computer when the S was performing incorrectly.

Essentially, two kinds of error cues were operative in this CAI course: Predictable errors or those that the author has anticipated from his teaching experience and the computer offers S specific remedial help; or Unpredictable errors that S will commit and therefore receives a generalization error cue. Table III lists the mean number of predictable cues and unpredictable cues* issued for treatment groups by chapters in the course.

* A student, or a pair, might conceivably receive a maximum of three to seven cues on any one problem.

These error rate data indicate that the effects of pairing High students has little effect on immediate performance on items as compared with their controls. This seems to hold both for predictable and unpredictable errors. With Low students, however, the results are somewhat different.

Table III

Mean Number of Predictable (p) and Unpredictable (u) Error Cues Issued for Treatment Groups by Chapters

		Chapter						
		2	3	4	5	6	7	8
High Pairs	p =	.88	2.9	2.3	1.4	4.5	6.5	.8
	u =	7.13	22.5	16.8	6.1	11.3	32.1	18.0
High Controls		1.1	2.8	2.2	1.0	5.0	5.8	1.4
		4.6	16.9	17.4	5.8	8.6	31.8	19.1
Low Pairs		.56	3.0	2.1	1.6	4.8	7.6	1.7
		4.4	18.2	15.6	10.7	11.4	37.3	18.8
Low Controls		1.3	3.2	2.2	1.7	5.5	9.8	1.8
		10.5	22.7	22.3	10.3	17.0	48.7	30.2

In chapter two, the error rate for Low pairs is less than half that of the Low controls. It is also noted that pairing Low students reduced unpredictable errors by approximately 25% in five of the remaining six chapters.

Attitudes

The questionnaire results are perhaps the most revealing of all since they indicate just how the person perceived himself and his partner in this study. Table IV lists the results for the four questions mentioned earlier in this paper.

It is noteworthy that 100% of the Ss in the Low Pairs rejected the pairs arrangement as compared with a rejection rate of 60% in the High

Pairs (question No. 5). A Fisher Exact Probability Test indicated this difference to be significant at less than .01 level of confidence. Apparently the Low Pairs were not necessarily rejecting their partners as persons since 60% would have chosen the same partner if forced to work in pairs again (question No. 6). The Fisher Test indicated that there was no reliable difference between High and Low Pairs in preference for the same or different partner.

In terms of the student's estimate of his own final exam performance, it is not so simple to state whether Highs or Lows as a group had the more realistic appraisal of their work. Apparently such a perception is a function of the pairing arrangement, CEEB score and the difficulty level of the exam. Low Pairs, for example, estimated their final scores on the average to be only 19% lower than they actually were, while the High Pairs downgraded themselves by as much as 33%. In the control group the trend was just reversed -- Low Controls judged themselves 33% below actual performance, while High Controls were 25% below (question No. 3).

The majority of the Ss volunteered information as to why they downgraded their performance on the questionnaire. Essentially, they felt ill at ease in this "new" no feedback situation with the paper and pencil test. Apparently they had grown too accustomed to learning statistics with such immediate feedback and intricate cues that they were experiencing withdrawal symptoms.

On the task of estimating whether the partner's final performance was higher or lower than his own (question No. 4), Low Pairs were correct only at chance level, 50%. High Pairs apparently were quite accurate in sizing up their partner since they were correct in 90% of the cases.

Table IV

Analysis of Questionnaire Results

Treatments	Questions			
	No. 3 ($\bar{X}$ estimate final)	No. 4 (estimate partner's final)	No. 5 (alone-partner)	No. 6 (same or different partner)
High Pairs	33.3% below actual	90% correct	60% alone 30% partner 10% no pref.	50% same 40% different 10% no pref.
High Controls	25.0% below actual			
Low Pairs	19.2% below actual	50% correct	100% alone	60% same 40% diff.
Low Controls	33.3% below actual			

DISCUSSION

The results from this study seem to indicate that students paired on CEEB verbal scores as a predictor variable will do as well as their controls on a final exam in a CAI course. In addition, they can complete the course in the same amount of time as their controls. The economic advantage, of course, is quickly realized since one has cut the cost of the educational terminal device by a factor of two in the process.

How widely these results can be generalized to other subject matters and predictor variables is a question open for further discussion and research. These results do agree closely, however, with Dick[2] in a study using pairs on a linear program in Algebra. No significant differences were found between pairs and controls on a final exam in that experiment. However, a retest on 80% of the Ss a year later yielded significantly greater retention scores in favor of the pairs. In the Dick study Ss were assigned to pairs at random so that further comparisons on aptitude or ability groupings are difficult if not impossible to make between these studies.

It would appear that further pairing studies on which students are matched on some predictor variable, including personality and attitude dimensions, would be a fruitful line of investigation. Attitudinal factors are argued for since it was apparent in this study that Low pairs completely rejected the pair arrangement, but not necessarily their partners.

In one way these results might seem at variance with a communication model of behavior which suggests that people desire to be with similar people in order to test the appropriateness of their response. However, if a machine can inform the Low S of the validity of his response, apparently he does not require a person to fulfill this role, preferring instead a non-threatening machine.

It is conjectured that another fruitful line of investigation in paired student instruction might be concerned with the parameters within the

final examination itself. While the final in this study was exclusively of the problem solving variety, a heavily weighted section on conceptual tasks might comprise a second examination. In other words, it is suggested here that in a group situation Ss might tend to become task oriented and, therefore, miss some of the conceptual framework of the course. Such deficiencies would tend to show up in such an exam of the kind proposed or in the acquisition of further material in that field.

It will be recalled from Table III that an analysis of the error rates by chapter indicated little difference between High pairs and their controls in this study but larger differences in instances between Low pairs and their controls. It was originally hypothesized that pairing students would reduce error rates in learning across both pair treatments because of the nature of the pair agreement rule. However, since pairing Low students will raise their immediate performance level on individual items, one would predict that long term retention would be improved for that group. This would square with an earlier study by Alter[1] in which she compared Ss retention curves from high, middle and low performers with respect to initial achievement, intelligence and time taken to read the program. Retention curves were plotted by retesting groups from the sample at differing time periods from initial learning. As a result, no significant differences were found between the retention curves of any of the subgroups when corrected for initial achievement with a covariance analysis. She concluded.

> These findings imply that if we are interested in improving retention we should operate primarily on improving the learner's initial achievement. This may be difficult with low I. Q. students. We may still expect differences in level of retention among the I. Q. groups, but this procedure should help to minimize these differences. These data give us no reason to believe that the lower I. Q. students will forget any more or less rapidly than the higher I. Q. students once they have been brought up to the same achievement level. (p. 6)

Perhaps, if for nothing more, the pairs arrangement has demonstrated a way of raising immediate performance levels for Low Ss without rewriting a program specifically for them. It will be for further studies to test the net result of this observation on tests of long term retention and the acquisition of extended material in the field.

References

1. Alter, Millicent, "Retention as a Function of Length of Retention Interval, Intelligence and Training Time", J. Programed Instruction, 1963, II (2), pp. 7-17.

2. Dick, Walter, "Retention as a Function of Paired and Individual Use of Programed Instruction," J. Programed Instruction, 1963 II (3), pp. 17-23.

3. Grubb, Ralph and Selfridge, Lenore, "The Computer Tutoring of Statistics," Computers and Automation, March 1964, pp. 20-26.

4. Uttal, William, "On Conversational Interaction," in Programed Learning and Computer Based Instruction, Coulson (ed.) New York: Wiley, 1962.